UNIVERSITAT ROVIRA I VIRGILI
ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA

Passing Network Analysis of the 2011 UCL Final

Master in Health Data Science — MHEDAS

Subject: Complex Networks

Author(s): Ravneet-Rahul Sandhu Singh,
Sofía González Estrada,
Teferi Mekonnen Yitayew

Date: January 14, 2026

Table of Contents

1. Introduction **1**
1.1 Goals 1
1.2 Objectives 1
1.3 Scope 2
2. Methodology **2**
2.1 Network Construction 2
2.1.1 Data Acquisition and Preprocessing 2
2.1.2 Node and Edge Definitions 3
2.2 Macro and Micro-Level Metrics 6
2.3 Community Detection 6
2.4 Percolation Analysis 6
2.5 Diffusion Analysis 6
References **6**

1. Introduction

The UEFA Champions League Final held on May 28, 2011, at Wembley Stadium is frequently cited as one of the most significant matches in modern football history. Pep Guardiola's FC Barcelona defeated Sir Alex Ferguson's Manchester United 3-1, but the scoreline only partially captures the nature of the event. It represented the culmination of the *Juego de Posición* a.k.a. *Tiki-Taka* philosophy, characterized by high passing frequency, fluid movement, and the tactical innovation of the *False 9*.

> *In my time as a manager, I would say they're the best team we've faced. Everyone acknowledges that and I accept that. It's not easy when you've been well beaten like that to think another way. No one has given us a hiding like that.*
>
> — ***Sir Alex Ferguson***, *Post-match press conference [1]*

1.1 Goals

Traditionally, football analysis has relied on aggregate statistics such as possession percentage, shots on target, or total passes. However, these metrics often fail to capture the structural complexity and the relational dynamics of a team. Football is inherently a system of interactions; a team can be modeled as a complex network where players are nodes and passes represent the edges linking them. This project applies Network Science to deconstruct the 2011 Final. By visualizing and analyzing the passing networks of both teams, we aim to uncover the underlying topology that allowed FC Barcelona to dominate Manchester United's defensive structure.

1.2 Objectives

The main goal of this project is to perform a comparative network analysis of FC Barcelona and Manchester United during the 2011 Champions League Final. To achieve this, we have established the following specific objectives:

- **Network metrics comparison:** Calculate and compare key network metrics (e.g., degree centrality, betweenness centrality, clustering coefficient) for both teams to identify structural differences in their passing networks. We aim to quantify the robustness of FC Barcelona's connectivity compared to the rigidity of United's structure.
- **Community structure analysis:** Analyze the community structures within each team's passing network to identify clusters of players who frequently interact. We expect FC Barcelona's network to exhibit more cohesive subgroups.
- **Role analysis:** Investigate the importance of individual players within the network by removing them and observing the impact on overall network connectivity and performance. Key players such as Messi, Pedro, and Busquets for FC Barcelona, and Rooney and Giggs for Manchester United will be the focus of this analysis.
- **Game build-up analysis:** Examine how effectively each team built up play from defense to attack through their passing networks.

1.3 Scope

The scope of this analysis is defined by the following constraints:

- **Data limitation:** The analysis is limited to this match only, no external data is considered which could provide additional context or comparative data.
- **Data representation:** The network is constructed solely from successful passing data. Off-the-ball movements, defensive actions (tackles, interceptions), and dribbles are not explicitly modeled as network edges or attributes.
- **Dynamic limitations:** While player positions may be visualized using average spatial coordinates to provide context, the core analysis focuses on topological metrics (graph properties) rather than spatial metrics.

2. Methodology

In this section, we outline the methodology employed to construct and analyze the passing networks of FC Barcelona and Manchester United.

2.1 Network Construction

2.1.1 Data Acquisition and Preprocessing

To construct the passing networks, we utilized event data obtained via the `statsbombpy` Python library [2], which interfaces with the StatsBomb Open Data API. We filtered the competition dataset for the 2010/2011 UEFA Champions League season to isolate the final match between FC Barcelona and Manchester United. The complete data acquisition and cleaning pipeline is illustrated in **Figure 2.1**.

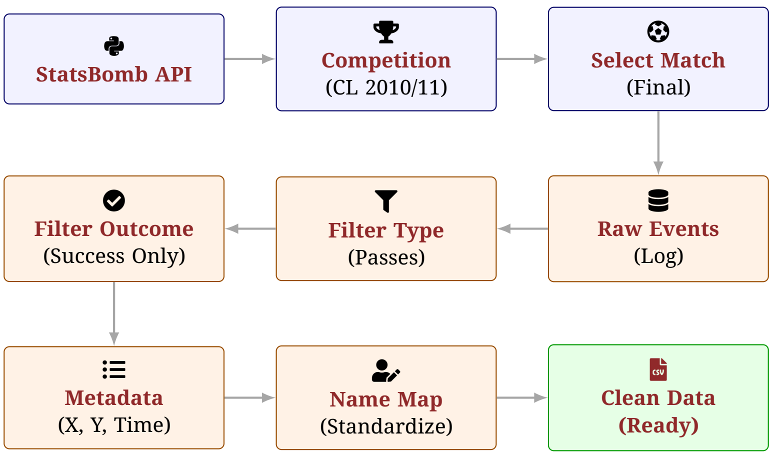

Figure 2.1: Data acquisition and preprocessing pipeline.

The extracted event log provides details for every ball action. For our analysis, we specifically filtered for events categorized as *Pass* where the `pass_outcome` attribute was null, indicating a successful completion. Each pass event includes crucial metadata used in the network generation: the identity of the passer and receiver, the timestamp (minute), and the (x, y) coordinates of the pass origin.

2.1.2 Node and Edge Definitions

The generated network is a directed, weighted graph $G = (V, E)$ where V represents the set of the top 11 players by completed pass count for each team, and E the set of edges weighted by pass frequency. Specifically:

- **Node Selection:** For each team, the 11 players with the highest number of completed passes were selected to ensure a fair comparison between teams and to focus on the primary contributors to each team's passing network.
- **Node Positioning:** The spatial position of each node was determined by calculating the centroid (mean x, y coordinates) of all successful passes initiated by that player during the observed time window.
- **Node Size:** Node markers were scaled proportionally to the player's total pass volume, with the zoom factor calculated as:

$$\text{Zoom} = 0.18 + \frac{\text{Pass Count}}{200} \tag{2.1}$$

- **Edge Attributes:** Edges were weighted by frequency. To enhance visual clarity, edge width and opacity were dynamically scaled based on the number of passes exchanged, with the edge width calculated as:

$$\text{Edge Width} = 0.5 + \frac{\text{Number of Passes}}{6} \tag{2.2}$$

  This means that for every 6 passes between two players, the edge width increases by 1 unit, starting from a minimum width of 0.5 units. Edge opacity was set to 0.4 for connections with fewer than 2 passes and 0.8 otherwise.

Figure 2.2 and **Figure 2.3** display the final networks that will be used for our analyses. Note that depending on the specific analysis, we may add or remove attributes from the edges, but the baseline network structure remains as shown. To enhance visualizations, we utilized player headshots from Transfermarkt [3] as node icons.

While in the beginning we considered temporal segmentation to capture the evolution of passing patterns, it also presents analytical challenges. On the one hand, for FC Barcelona, the late substitution windows contain sparse data due to a more defensive mindset to conserve the lead in the final minutes. On the other hand, for Manchester United, the shorter temporal segments result in insufficient edge density to identify structural patterns.

To address these limitations and obtain a more comprehensive view of each team's passing behavior, we constructed full match networks (0'–94') that aggregate all passing interactions across the entire game. Importantly, these networks include all players who participated, with substitutes appearing as separate nodes rather than merging their contributions with the players they replaced. This approach provides several advantages: it maintains the complete picture of each team's squad, preserves the individual contributions of substitute players, and generates sufficient data density for robust network analysis.

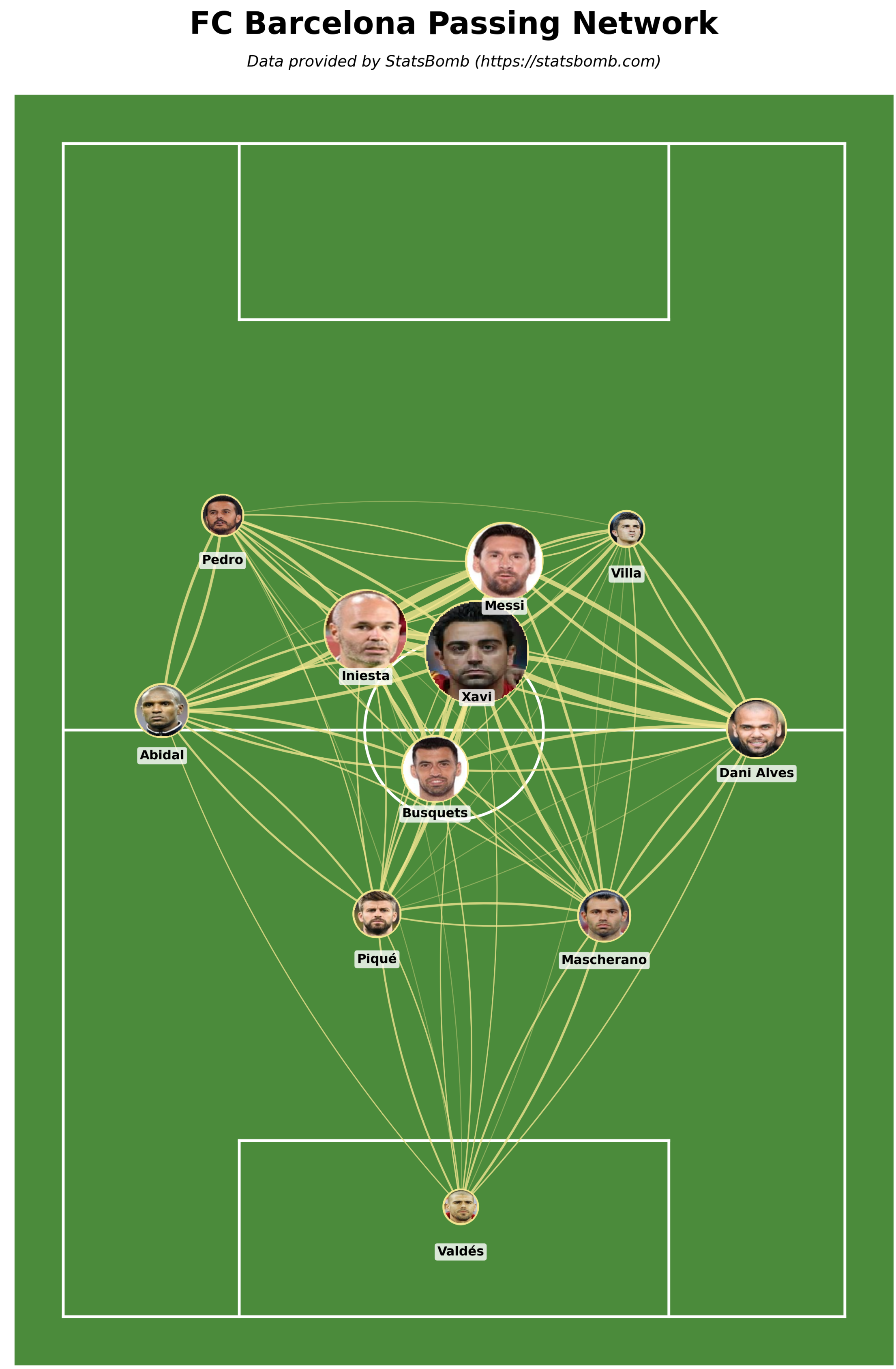

Figure 2.2: FC Barcelona full match passing network.

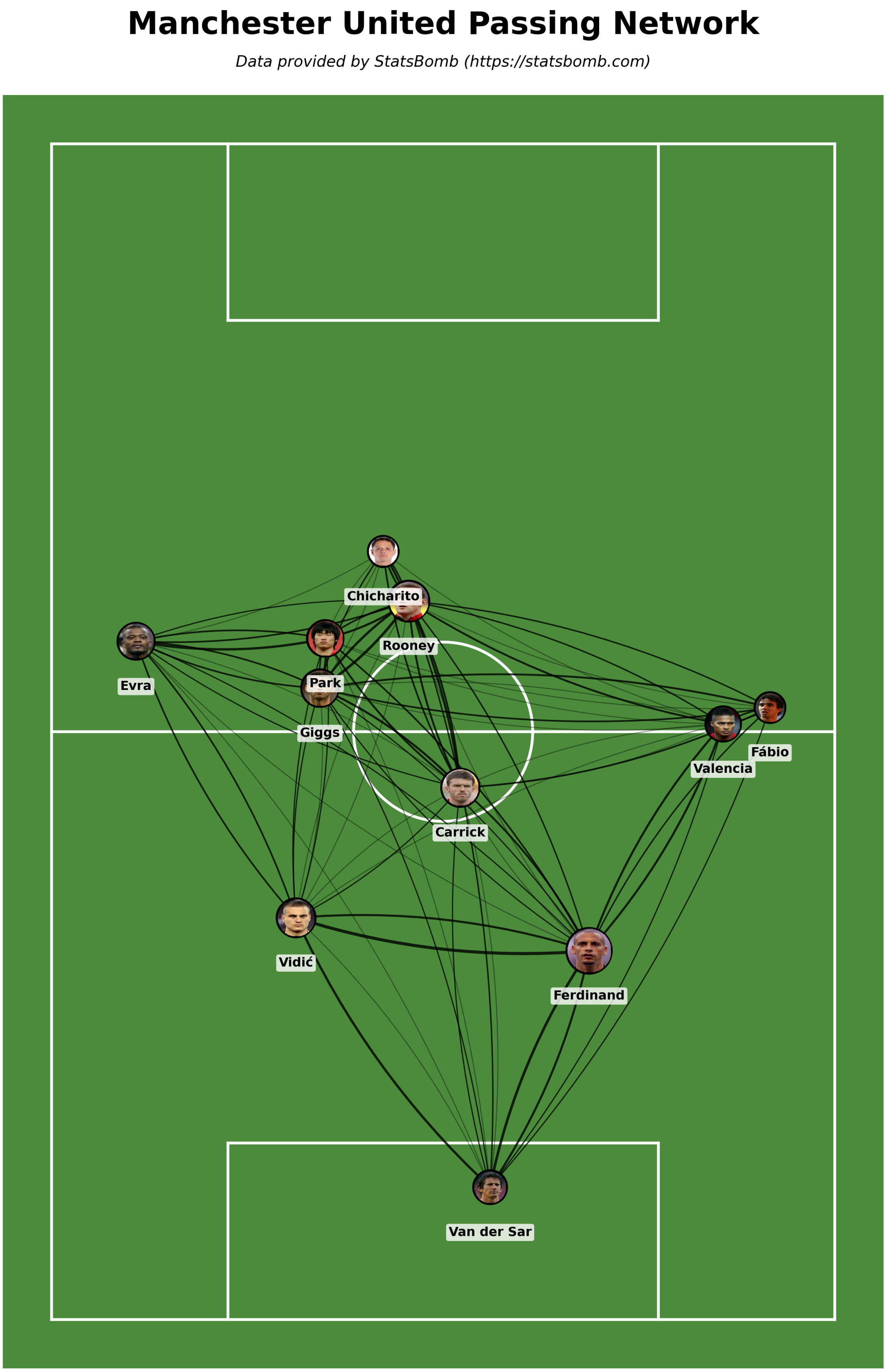

Figure 2.3: Manchester United full match passing network.

2.2 Macro and Micro-Level Metrics

We computed macro-level (global) and micro-level (node-specific) metrics to quantitatively characterize the passing networks of both teams. The network was constructed as a directed weighted graph, where edges represent passes from sender to receiver and weights (w_{ij}) indicate pass frequency. To accurately model the efficiency of ball circulation for path-based metrics (Shortest Path and Betweenness), we defined the distance or cost of an edge as the inverse of its frequency, $d_{ij} = 1/w_{ij}$.

At the macro-level, we calculated:

- **Average Degree:** Measures the average level of player involvement in the game. A higher value indicates a team that dominates possession through a high volume of passes, rather than relying on individual runs or direct play.
- **Average Clustering Coefficient:** Quantifies the tendency of players to form passing triangles. A higher value indicates a style relying on tight, short combinations and triangular structures, whereas lower values suggest a more linear or direct playstyle.
- **Assortativity:** Assesses whether high-volume passers tend to connect with other high-volume passers. A positive value implies a key players pass primarily among themselves, while a negative value indicates a hierarchical structure where playmakers distribute the ball to less central players.
- **Average Shortest Path Length:** Reflects the efficiency and speed of ball circulation across the team. A lower value indicates less resistance in the network, implying that the ball can move between any two players rapidly through frequently used passing lanes.

At the micro-level, we computed:

- **Degree Centrality:** Measures the number of unique teammates a player connects with relative to the team size. High values highlight players who act as broad distributors, linking with a wide variety of teammates across the pitch rather than focusing on a single partner.
- **Betweenness Centrality:** Identifies players who act as bridges along the most efficient passing routes. High values indicate players who are critical for transition, connecting otherwise disconnected areas of the field (e.g., linking defense to attack).
- **Eigenvector Centrality:** Measures a player's influence based on the importance of the teammates they connect with. High values identify players who are key in the overall network structure, frequently exchanging the ball with other influential playmakers.

2.3 Community Detection

2.4 Percolation Analysis

2.5 Diffusion Analysis

References

[1] D. Fifield, "Sir Alex Ferguson: No one has given United a hiding like Barcelona did," *The Guardian*, May 2011. Accessed: 2025-12-23.

[2] StatsBomb, "statsbombpy: A Python package to easily stream StatsBomb data." https://github.com/statsbomb/statsbombpy, 2025. Maintained by Hudl. Accessed: 2025-12-23.

[3] Transfermarkt, "Football transfers, rumours, market values and news," 2026. Accessed: January 5, 2026.